Wirtschaftswissenschaftliches Prüfungsamt

Bachelor-Prüfung
Schließende Statistik
Sommersemester 2013

Aufgabenstellung und Ergebnisse

Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Kleben Sie bitte **sofort** Ihr Namensschild auf obige Markierung!
- Die Klausur besteht aus insgesamt 8 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 16 + 12 + 8 + 16 + 18 + 16 + 8 + 26) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbsterstellte DIN-A4 Blätter bzw. 4 selbsterstellte (einseitige) DIN-A4 Seiten. Benötigte Tabellen zur Normal- und t-Verteilung finden Sie am Ende dieses Klausurheftes.
- Bei der Korrektur werden **nur** die Lösungen auf den Seiten 1–10 berücksichtigt. Das letzte Blatt (Tabellen zur Normal- und t-Verteilung) darf abgetrennt werden.
- Bei mehreren Lösungsvorschlägen muss die gültige Lösung eindeutig gekennzeichnet sein.
- Mit Ausnahme der Multiple-Choice-Aufgaben muss der Lösungsweg klar ersichtlich sein.

Bewertungsteil — Bitte nicht beschreiben								
Aufgabe	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Σ
1								
2								
3								
4								
5								
6								
7								
8								
Σ								

Aufgabe 1 (16 Punkte)

Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben +2 Punkte, falsche Antworten −1 Punkt, nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

		wahr	falsch
1.	Für die Erwartungswerte der Zufallsvariablen $X_1, \ldots, X_n$ und Y gelte $\mathrm{E}(X_1) = \ldots = \mathrm{E}(X_n) = \mathrm{E}(Y)$ sowie für die Varianzen $\mathrm{Var}(X_1) = \ldots = \mathrm{Var}(X_n) = \mathrm{Var}(Y)$. Dann ist $X_1, \ldots, X_n$ stets eine einfache Stichprobe zu Y.	☐	■
2.	Sind für $n \in \mathbb{N}$ Schätzfunktionen T_n gegeben mit $\lim_{n\to\infty} \mathrm{Var}(T_n) = \lambda$, dann ist die Familie von Schätzfunktionen T_n stets effizient für λ.	☐	■
3.	Es sei $X_1, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zur Zufallsvariablen Y (mit existierendem Erwartungswert und existierender Varianz). Dann gilt für den Stichprobenmittelwert $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ stets $\mathrm{E}(\overline{X}) = \mathrm{E}(Y)$ und $\mathrm{Var}(\overline{X}) = \dfrac{\mathrm{Var}(Y)}{n}$.	■	☐
4.	Die Breite von (approximativen) Konfidenzintervallen für den Parameter p einer alternativverteilten Zufallsvariablen hängt (über den Schätzwert für p) jeweils von der konkreten Stichprobenrealisation ab.	■	☐
5.	Bei einem linksseitigen Gauß-Test von $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$ kann die Fehlerwahrscheinlichkeit 2. Art für $\mu < \mu_0$ durch $1 - G(\mu)$ berechnet werden, wobei $G(\mu)$ die Gütefunktion des Tests bezeichnet.	■	☐
6.	Für $\alpha \in (0,1)$ gilt für die Quantile der Standardnormalverteilung der Zusammenhang $N_\alpha = -N_{1-\alpha}$.	■	☐
7.	Beim Chi-Quadrat-Test für die Varianz sind nur zweiseitige Hypothesentests möglich.	☐	■
8.	Bei der Anwendung des Chi-Quadrat-Anpassungstests ist die Annahme, dass die Grundgesamtheit Y normalverteilt ist, wesentlich. Kann diese Annahme nicht getroffen werden, ist auch eine näherungsweise Verwendung des Tests nicht mehr angebracht.	☐	■

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben +3 Punkte, falsche Antworten −1 Punkt, nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

1. Seien $Y \sim N(\mu, \sigma^2)$ und $X_1, \dots, X_n$ eine einfache Stichprobe vom Umfang n zu Y. Dann gilt mit $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ und $S := \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$:

   (a) $\frac{\overline{X}-\mu}{S^2}\sqrt{n} \sim t(n-1)$ ☐

   (b) $\frac{\overline{X}-\mu}{S}\sqrt{n} \sim t(n-1)$ ■

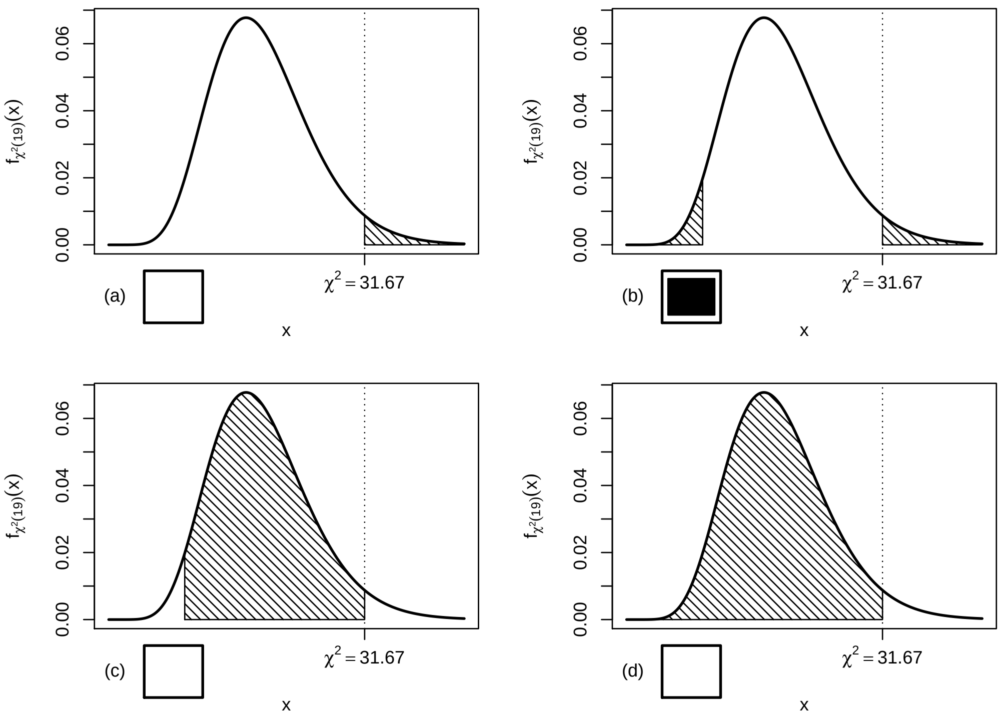
   (c) $\frac{\overline{X}-\mu}{S^2}\sqrt{n} \sim t(n)$ ☐

   (d) $\frac{\overline{X}-\mu}{S}\sqrt{n} \sim t(n)$ ☐

2. Sei $X_1, \dots, X_{20}$ eine einfache Stichprobe zu einer $N(\mu, \sigma^2)$-verteilten Zufallsvariablen Y mit unbekannten Parametern μ und σ^2. Auf der Grundlage einer Stichprobenrealisation zu dieser einfachen Stichprobe vom Umfang $n = 20$ soll

$$H_0 : \sigma^2 = \sigma_0^2 = 10 \qquad \text{gegen} \qquad H_1 : \sigma^2 \neq \sigma_0^2 = 10$$

   mit einem Chi-Quadrat-Test getestet werden. Als realisierte Teststatistik erhält man $\chi^2 = 31.67$. Markieren Sie die Abbildung, welche den p-Wert in der beschriebenen Situation korrekt als Inhalt der schraffierten Fläche unter der Dichtefunktion der Verteilung der Teststatistik unter H_0 (für $\sigma^2 = \sigma_0^2$) darstellt.

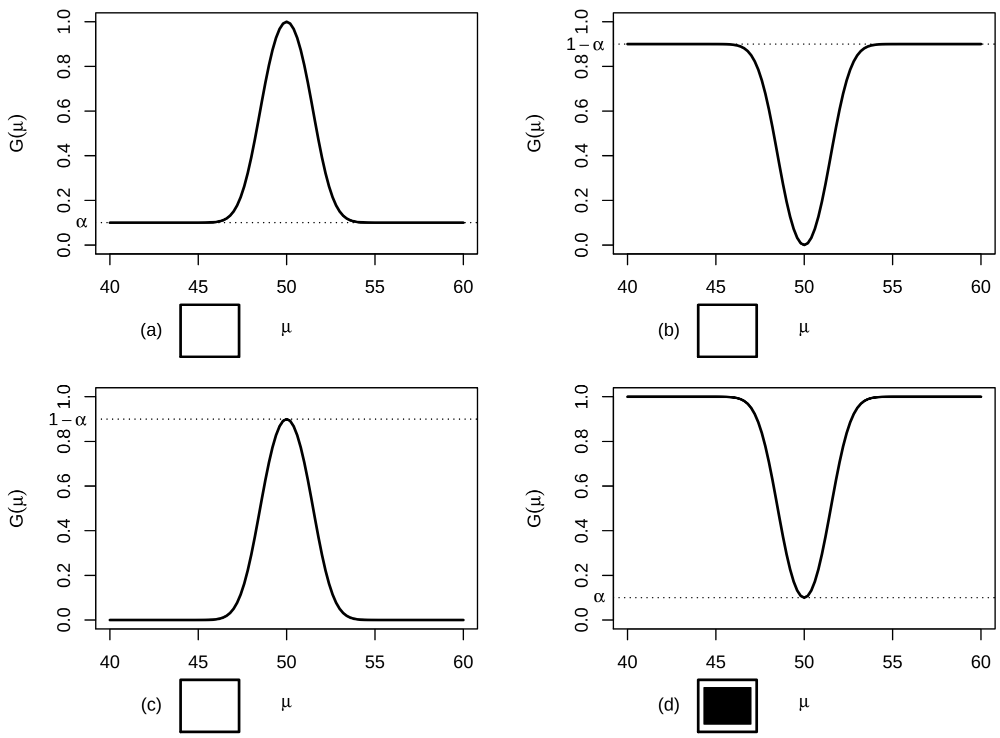


3. Als p-Wert zur realisierten Teststatistik eines zweiseitigen Gauß-Tests für den Mittelwert einer normalverteilten Grundgesamtheit bei bekannter Varianz erhält man $p = 0.01547$. Dann gilt:
   - (a) Die Nullhypothese ist bei einem Test zum Niveau $\alpha = 0.05$ abzulehnen, bei einem Test zum Niveau $\alpha = 0.025$ jedoch nicht abzulehnen. ☐
   - (b) Die Nullhypothese ist bei einem Test zum Niveau $\alpha = 0.025$ abzulehnen, bei einem Test zum Niveau $\alpha = 0.05$ jedoch nicht abzulehnen. ☐
   - (c) Die Nullhypothese ist sowohl bei einem Test zum Niveau $\alpha = 0.025$ als auch bei einem Test zum Niveau $\alpha = 0.05$ abzulehnen ■
   - (d) Die Nullhypothese ist weder bei einem Test zum Niveau $\alpha = 0.025$ noch bei einem Test zum Niveau $\alpha = 0.05$ abzulehnen. ☐

4. Auf der Grundlage einer einfachen Stichprobe $X_1, \ldots, X_{30}$ vom Umfang $n = 30$ zu einer $N(\mu, 5^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

$$H_0 : \mu = 50 \qquad \text{gegen} \qquad H_1 : \mu \neq 50$$

bei einem Signifikanzniveau von $\alpha = 0.1$ betrachtet.
Markieren Sie die Abbildung, welche die Gütefunktion des oben genannten Tests korrekt darstellt.



Aufgabe 3 (6 + 2 = 8 Punkte)

Die Verteilung einer Zufallsvariablen Y sei in Abhängigkeit des unbekannten Parameters $a > 0$ durch die folgende Dichtefunktion gegeben:

$$f_Y(y) = \begin{cases} a^2 \cdot (y-1) \cdot e^{-a \cdot (y-1)} & \text{für } y > 1 \\ 0 & \text{sonst} \end{cases}$$

Der Parameter a soll auf Grundlage einer einfachen Stichprobe $X_1, \ldots, X_n$ vom Umfang n geschätzt werden.

(a) Bestimmen Sie den Schätzer $\widehat{a}_{ML}$ nach der Maximum-Likelihood-Methode.

(b) Man kann zeigen, dass $\mathrm{E}(Y) = \frac{2}{a} + 1$ gilt. Bestimmen Sie damit den Schätzer $\widehat{a}_{MM}$ nach der Methode der Momente.

Hinweise:

- *Beachten Sie, dass Sie Teil (b) auch ohne die Bearbeitung von Teil (a) lösen können.*
- *Falls sich der ML-Schätzer als lokale Extremstelle einer differenzierbaren Funktion bestimmen lässt, muss* **nicht** *überprüft werden (z.B. mit Hilfe der 2. Ableitung), ob tatsächlich eine Maximalstelle vorliegt.*

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{a}_{ML} = \dfrac{2 \cdot n}{\sum_{i=1}^{n} (x_i - 1)} = \dfrac{2}{\overline{x} - 1}$

(b) $\widehat{a}_{MM} = \dfrac{2}{\overline{x} - 1}$

Aufgabe 4 (7 + 2 + 4 + 3 = 16 Punkte)

Bei der Abfüllung von Tafelwasser weiß der Hersteller aus langjähriger Erfahrung, dass die verwendete Maschine eine Standardabweichung von $5[ml]$ für die abgefüllte Menge hat. Nach einer routinemäßigen Überprüfung hat der Hersteller den Verdacht, dass die Abfüllanlage im Mittel mehr als die auf dem Produkt ausgezeichneten $750[ml]$ in die Flaschen einfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 36 Flaschen entnommen, deren gemessene Füllmengen $x_1, \ldots, x_{36}$ als Realisation einer einfachen Stichprobe vom Umfang 36 zur annahmegemäß $N(\mu, 5^2[ml^2])$-verteilten Abfüllmenge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{36} \sum_{i=1}^{36} x_i = 751.939[ml] \ .$$

(a) Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob der Verdacht der Herstellerfirma bestätigt werden kann. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

(b) Berechnen Sie den p-Wert zum Test aus Teil (a).

(c) Wie groß ist die Wahrscheinlichkeit des Fehlers 2. Art, falls $\mu = 751[ml]$ beträgt?

(d) Geben Sie auf Basis der Stichprobenrealisation ein zweiseitiges Konfidenzintervall für die mittlere Abfüllmenge zur Sicherheitswahrscheinlichkeit $1 - \alpha = 0.98$(!) an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $N = 2.327 \in (1.645, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test bestätigt also den Verdacht der Herstellerfirma, dass die von der Maschine abgefüllte Menge im Mittel zu hoch ist.

(b) $p = 0.0099$

(c) $\beta(751) = 0.67$

(d) Realisiertes symm. Konfidenzintervall zum Konf.-niveau $1-\alpha = 0.98$: $[750.001, 753.877]$

Aufgabe 5 (10 + 8 = 18 Punkte)

Es werde angenommen, dass die in Bar gemessene maximale Druckfestigkeit von Hochdruckschläuchen (bis zur Undichtigkeit) Y^A eines aktuell in der Produktion eines Unternehmens verwendeten Fabrikats normalverteilt sei mit unbekanntem Erwartungswert μ_A und unbekannter Varianz σ_A^2. Es wird erwogen, in der Produktion zukünftig ein alternatives Fabrikat zu verwenden, dessen maximale Druckfestigkeit Y^B ebenfalls als normalverteilt (mit unbekanntem Erwartungswert μ_B und unbekannter Varianz σ_B^2) angenommen werden kann. Es soll überprüft werden, ob das alternative Fabrikat im Mittel eine höhere maximale Druckfestigkeit als das aktuell verwendete Fabrikat besitzt.
Aus einer Druckprüfung von $n_A = 15$ Schläuchen des aktuell verwendeten und $n_B = 12$ Schläuchen des alternativen Fabrikats erhielt man Realisationen jeweils voneinander unabhängiger einfacher Stichproben $X_1^A, \ldots, X_{15}^A$ zu Y^A sowie $X_1^B, \ldots, X_{12}^B$ zu Y^B und hieraus die zugehörigen Mittelwerte $\overline{x^A} = 702.28$ bzw. $\overline{x^B} = 714.35$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 419.35$ bzw. $s_{Y^B}^2 = 332.11$.

(a) Testen Sie unter der Annahme $\sigma_A^2 = \sigma_B^2$ zum Signifikanzniveau $\alpha = 0.05$ die Hypothese, dass das alternative Fabrikat im Mittel eine höhere maximale Druckfestigkeit als das aktuell verwendete Fabrikat besitzt. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

(b) Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.1$, ob die in Teil (a) getroffene Annahme der Varianzgleichheit auf Grundlage der vorhandenen Stichprobeninformation verworfen werden muss. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

Hinweis: Verwenden Sie für Teil (b) den folgenden Tabellenausschnitt mit 0.95*-Quantilen von* $F(m,n)$*-Verteilungen sowie ggf. die Rechenregel* $F_{m,n;p} = \frac{1}{F_{n,m;1-p}}$.

$n \backslash m$	11	12	13	14	15	16	17	18	19	20
11	2.818	2.788	2.761	2.739	2.719	2.701	2.685	2.671	2.658	2.646
12	2.717	2.687	2.660	2.637	2.617	2.599	2.583	2.568	2.555	2.544
13	2.635	2.604	2.577	2.554	2.533	2.515	2.499	2.484	2.471	2.459
14	2.565	2.534	2.507	2.484	2.463	2.445	2.428	2.413	2.400	2.388
15	2.507	2.475	2.448	2.424	2.403	2.385	2.368	2.353	2.340	2.328
16	2.456	2.425	2.397	2.373	2.352	2.333	2.317	2.302	2.288	2.276
17	2.413	2.381	2.353	2.329	2.308	2.289	2.272	2.257	2.243	2.230
18	2.374	2.342	2.314	2.290	2.269	2.250	2.233	2.217	2.203	2.191
19	2.340	2.308	2.280	2.256	2.234	2.215	2.198	2.182	2.168	2.155
20	2.310	2.278	2.250	2.225	2.203	2.184	2.167	2.151	2.137	2.124

Ergebnisse (ohne Begründung/Rechenweg):

(a) $t = -1.597 \notin (-\infty, -1.708) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Der Test kann also die Vermutung, dass das alternative Fabrikat im Mittel eine höhere maximale Druckfestigkeit als das aktuell verwendete Fabrikat besitzt, nicht bestätigen.

(b) $F = 1.263 \notin [0, 0.39) \cup (2.739, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Der Test findet also keine Anzeichen für eine Verletzung der in Teil (a) angenommenen Varianzgleichheit.

Aufgabe 6 (16 Punkte)

Um zu überprüfen, ob es einen Zusammenhang zwischen der Anzahl bearbeiteter Zusatzübungsblätter und dem Abschneiden in der Klausur (bestanden/nicht bestanden) gibt, hat der Dozent einer Statistik-Veranstaltung aus den Korrekturergebnissen der zugehörigen Klausuren aller 292 Teilnehmer die folgende Tabelle zusammengestellt:

	0 Blätter bearbeitet	1 Blatt bearbeitet	2 Blätter bearbeitet
bestanden	104	45	91
nicht bestanden	42	8	2

Überprüfen Sie anhand dieses Datenmaterials zum Signifikanzniveau $\alpha = 0.01$ **(!)**, ob die Anzahl bearbeiteter Zusatzübungsblätter und das Klausurergebnis stochastisch unabhängig sind.

Hinweis: Verwenden Sie den folgenden Tabellenausschnitt mit p-Quantilen von $\chi^2(n)$-Verteilungen:

$n\backslash p$	0.01	0.025	0.05	0.5	0.90	0.95	0.975	0.99
1	0.000	0.001	0.004	0.455	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	1.386	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	2.366	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	3.357	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	4.351	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	5.348	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	6.346	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	7.344	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	8.343	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	9.342	15.987	18.307	20.483	23.209

Ergebnisse (ohne Begründung/Rechenweg):

$\chi^2 = 27.823 \in (9.21, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Der Test kommt also zum Ergebnis, dass die Anzahl bearbeiteter Zusatzübungsblätter und das Klausurergebnis nicht stochastisch unabhängig sind.

Aufgabe 7 (1 + 1 + 1 + 1 + 1 + 2 + 1 = 8 Punkte)

Zur Erklärung der stetigen Monatsrenditen der Allianz-Aktie y_i (in Prozent) durch die stetigen Monatsrenditen des DAX x_i (in Prozent) unterstellt man die Gültigkeit eines Zusammenhangs im Sinne des folgenden linearen Modells:

$$y_i = \beta_1 + \beta_2 x_i + u_i \qquad \text{mit} \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2),\ i \in \{1, \ldots, n\}$$

Aus Daten der XETRA-Börse der jüngeren Vergangenheit wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

```
Call:
lm(formula = y ~ x)

Residuals:
   Min     1Q Median     3Q    Max
-7.537 -1.712 -0.377  1.353 10.163

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.148      1.885  -0.609   0.5595
x              1.781      0.556   3.203   0.0126 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.159 on 8 degrees of freedom
Multiple R-squared:  0.5618,       Adjusted R-squared:  0.5071
F-statistic: 10.26 on 1 and 8 DF,  p-value: 0.01256
```

(a) Wie viele Monatsrenditen gingen in die Schätzung ein?

(b) Geben Sie die realisierten Kleinst-Quadrate-Schätzwerte für β_1 und β_2 an.

(c) Geben Sie den realisierten Schätzwert für σ^2 an.

(d) Welcher Anteil der Gesamtvarianz der stetigen Monatsrenditen der Allianz-Aktie wird durch das lineare Modell erklärt?

(e) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.01$, ob β_2 signifikant positiv ist.

(f) Geben Sie ein Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für β_1 an.

(g) Welche stetige Monatsrendite der Allianz-Aktie prognostiziert das Modell in einem Monat mit stetiger DAX-Rendite von 1.35 (in Prozent)?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $n = 10$

(b) $\widehat{\beta}_1 = -1.148$, $\widehat{\beta}_2 = 1.781$

(c) $\widehat{\sigma^2} = 26.6153$

(d) 0.5618

(e) β_2 ist signifikant positiv.

(f) $[-5.4948, 3.1988]$

(g) $\widehat{y}_0 = 1.2564$

Aufgabe 8 (6 + 2 + 2 + 3 + 5 + 3 + 5 = 26 Punkte)

Zur Schätzung eines einfachen linearen Regressionsmodells

$$y_i = \beta_1 + \beta_2 \cdot x_i + u_i \qquad \text{mit} \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2), \quad i \in \{1, \ldots, n\}$$

aus einer Stichprobe vom Umfang $n = 30$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{30} y_i = 658.902; \quad \sum_{i=1}^{30} y_i^2 = 14806.085; \quad \sum_{i=1}^{30} x_i = 180.998;$$

$$\sum_{i=1}^{30} x_i^2 = 1175.961; \quad \sum_{i=1}^{30} x_i \cdot y_i = 4088.142; \quad \sum_{i=1}^{30} \widehat{y}_i^2 = 14623.331;$$

(a) Schätzen Sie β_1 und β_2 mit Hilfe der Kleinst-Quadrate-Methode.

(b) Berechnen Sie das Bestimmtheitsmaß R^2.

(c) Geben Sie mit Hilfe der bekannten erwartungstreuen Schätzfunktion für σ^2 den realisierten Schätzwert für σ^2 an.

(d) Berechnen Sie $\widehat{\sigma^2}_{\widehat{\beta}_1}$ und $\widehat{\sigma^2}_{\widehat{\beta}_2}$.

(e) Überprüfen Sie zum Signifikanzniveau $\alpha = 0.01$ **(!)**, ob β_2 signifikant positiv ist. Fassen Sie das Ergebnis auch in einem Antwortsatz zusammen.

(f) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für β_1 an.

(g) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für y_0 gegeben $x_0 = 5$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 13.849$, $\widehat{\beta}_2 = 1.345$

(b) $R^2 = 0.4539$

(c) $\widehat{\sigma}^2 = 6.53$

(d) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 3.045$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.07768$

(e) $t = 4.826 \in (2.467, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

β_2 ist also signifikant positiv.

(f) $[10.275, 17.423]$

(g) $[15.221, 25.927]$